

REPORT



과목명		비모수통계자료분석
담당교수		최영훈 교수님
학과		응용통계학과
학년		3학년
학번		201452024
이름		박상희
제출일		2018년 9월 17일

연습문제 1

다음은 금연교육 프로그램의 실시 전과 실시 후의 체중비교 자료이다(단위:kg). 부호검정을 이용하여 금연 후의 체중이 감소하는지에 대한 p값을 구하여라.

금연 전	66	80	69	52	75
금연 후	71	82	68	50	73

(1) 쌍비교

X	66	80	69	52	75
Y	71	82	68	50	73
쌍비교	+	+	-	-	-

(2) 가설

$$H_0 : P(+) \geq P(-)$$

$$H_1 : P(+) < P(-)$$

혹은

$$H_0 : E(X) \geq E(Y)$$

$$H_1 : E(X) < E(Y)$$

혹은

$$H_0 : \text{체중변화가 없다.}$$

$$H_1 : \text{체중이 감소한다.}$$

(3) 검정통계량

T = “+” 쌍의 총 개수 = 2

(4) 유의수준

$$\textcircled{1} \quad T = 2 \sim B(5, \frac{1}{2})$$

$P(T \leq 2 \mid E(T) = np = 2.5)$ 이 값을 이항분포표에서 찾으면.... $P = 0.5000$ 이다.

$$\textcircled{2} \quad P(T \leq 2) = P\left(\frac{T - E(T)}{\sqrt{\text{Var}(T)}} \leq \frac{2 - np}{\sqrt{npq}}\right) = P\left(Z \leq \frac{2 - 2.5}{\sqrt{1.25}}\right) = P(Z \leq -0.4472) = 0.3275$$

유의수준 5% 하에서 “체중이 감소하지 않는다.” 라는 귀무가설을 기각할만한 근거가 충분하지 않다. 따라서 금연 프로그램으로 인한 체중의 감소는 없다고 볼 수 있다.

연습문제 2

감기약을 복용하면 혈압이 떨어지고 현기증을 일으키는 것으로 알려져 있다. 이를 확인하기 위하여 20명의 감기 환자를 무작위 추출하여 감기약 복용 30분 전과 3시간 후의 혈압을 기록하였다. 가설을 설정하고 부호검정을 이용하여 검정하여라.

복용 전	90	76	69	62	55	71	54	64	72	69	73	82	76	77	74	89	75	69	70	92
복용 후	72	74	76	55	58	59	50	61	55	60	70	66	77	62	59	74	73	67	66	82

(1) 쌍비교

X	90	76	69	62	55	71	54	64	72	69	73	82	76	77	74	89	75	69	70	92
Y	72	74	76	55	58	59	50	61	55	60	70	66	77	62	59	74	73	67	66	82
쌍비교	-	-	+	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-

(2) 가설

$$H_0 : P(+) \geq P(-)$$

$$H_1 : P(+) < P(-)$$

혹은

$$H_0 : E(X) \geq E(Y)$$

$$H_1 : E(X) < E(Y)$$

혹은

$$H_0 : \text{혈압변화가 없다.}$$

$$H_1 : \text{혈압이 감소한다.}$$

(3) 검정통계량

T = “+” 쌍의 총 개수 = 3

(4) 유의수준

$$\textcircled{1} \quad T = 3 \sim B(20, \frac{1}{2})$$

$P(T \leq 3 \mid E(T) = np = 10)$ 이 값을 이항분포표에서 찾으면.... $P = 0.0013$ 이다.

$$\textcircled{2} \quad P(T \leq 3) = P\left(\frac{T - E(T)}{\sqrt{\text{Var}(T)}} \leq \frac{3 - np}{\sqrt{npq}}\right) = P\left(Z \leq \frac{3 - 10}{\sqrt{5}}\right) = P(Z \leq -3.1304) = 0.0008$$

유의수준 5% 하에서 “혈압이 감소하지 않았다.” 라는 귀무가설을 기각할만한 근거가 충분하다. 따라서 감기약으로 인해 혈압이 떨어졌다고 볼 수 있다.

연습문제 3

두 대통령후보 A와 B 간의 TV 토론이 유권자에게 영향력을 미치는지를 알아보기 위하여 임의로 추출한 25명 유권자의 두 대통령후보에 대한 TV 토론 전후 선호도의 여론조사를 실시한 자료가 아래와 같았다. McNemar 검정을 실시하여라.

		TV토론 후		
		A	B	
TV토론 전	A	10	26	36
	B	10	9	16
		20	35	55

(1) 쌍비교

	$Y_i = 0$	$Y_i = 1$
$X_i = 0$	10	26 (+)
$X_i = 1$	10 (-)	9

(2) 가설

$$H_0 : P(+) = P(-)$$

$$H_1 : P(+) \neq P(-)$$

혹은

$$H_0 : E(X) = E(Y)$$

$$H_1 : E(X) \neq E(Y)$$

혹은

$$H_0 : \text{지지변화가 없다.}$$

$$H_1 : \text{지지변화가 있다.}$$

(3) 검정통계량

T = “+” 쌍의 총 개수 = 26

(4) 유의수준

$$\textcircled{1} \quad T = 26 \sim B(36, \frac{1}{2}) \quad P(T \geq 26 \mid E(T) = np = 18) \quad n = 25 \text{ 이상이면 찾기 힘들다...}$$

$$\textcircled{2} \quad 2 \times P(T \geq 26) = 2 \times P\left(\frac{T - E(T)}{\sqrt{\text{Var}(T)}} \geq \frac{26 - np}{\sqrt{npq}}\right) = 2 \times P\left(Z \geq \frac{26 - 18}{\sqrt{9}}\right)$$

$$= 2 \times P(Z \geq 2.67) = 2 \times [1 - P(Z \leq 2.665)] = 2 \times [1 - 0.996] = 0.008$$

$$T^* = \frac{T - E(T)}{\sqrt{\text{Var}(T)}} = \frac{b - np}{\sqrt{npq}} = \frac{b - \frac{n}{2}}{\frac{\sqrt{n}}{2}} = \frac{2b - n}{\sqrt{n}} = \frac{b - c}{\sqrt{b + c}} = \frac{26 - 10}{\sqrt{26 + 10}} = \frac{16}{6} = 2.67 \sim N(0, 1)$$

$$\textcircled{3} \quad T^* = \left(\frac{b - c}{\sqrt{b + c}}\right)^2 = \frac{(b - c)^2}{b + c} \sim \chi_1^2$$

$$T^* = 7.29 \sim \chi_1^2 \quad \hat{\alpha} = P(T^* \geq 7.29) = 1 - P(T^* \leq 7.29) = 1 - 0.995 = 0.005$$

유의수준 5% 하에서 “유권자들의 지지변화가 없다.” 라는 귀무가설을 기각할만한 근거가 충분하다. 따라서 TV 토론으로 인해 유권자들의 지지변화가 생겼다고 볼 수 있다.

연습문제 4

자동차 보험회사에서는 자동차 사고로 인한 보상 보험금액이 지속적으로 증가하고 있다고 주장한다. 이를 확인하기 위하여 보험회사 경영진은 오래된 시간 순으로 임의 추출한 15건의 자동차 사고로 인하여 지급된 보험금액 자료(단위:만원)를 다음과 같이 얻을 수 있다. 유의수준 5% 하에서 필요한 검정을 하여라.

72.7 21.5 122.7 35.9 27.4 32.6 65.4 47.5 52.8 75.1
67.5 99.7 135.2 183.5 55.3

(1) 쌍비교

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
72.7	21.5	122.7	35.9	27.4	32.6	65.4	47.5	52.8	75.1	67.5	99.7	135.2	183.5	55.3

순서쌍	실제값	쌍비교
(x_1, x_9)	(72.7, 52.8)	-
(x_2, x_{10})	(21.5, 75.1)	+
(x_3, x_{11})	(122.7, 67.5)	-
(x_4, x_{12})	(35.9, 99.7)	+
(x_5, x_{13})	(27.4, 135.2)	+
(x_6, x_{14})	(32.6, 183.5)	+
(x_7, x_{15})	(65.4, 55.3)	-

(2) 가설

$H_0 : P(+) \leq P(-)$ $H_0 : E(X) \leq E(Y)$ $H_0 : \text{보상 보험금액이 증가하지 않았다.}$
 $H_1 : P(+) > P(-)$ $H_1 : E(X) > E(Y)$ $H_1 : \text{보상 보험금액이 증가하였다.}$

(3) 검정통계량

T = “+” 쌍의 총 개수 = 4

(4) 유의수준

$$\textcircled{1} \quad T = 4 \sim B(7, \frac{1}{2}) \quad P(T \geq 4 \mid E(T) = np = 3.5) = 1 - P(T \leq 3)$$

이 값을 이항분포표에서 찾으려면.... $1 - P = 1 - 0.5 = 0.5$ 이다.

$$\textcircled{2} \quad T^* = \frac{b - c}{\sqrt{b + c}} = \frac{4.5 - 3}{\sqrt{4 + 3}} = \frac{1.5}{\sqrt{7}} = 0.5669 \sim N(0, 1)$$

$$\hat{\alpha} = P(Z \geq 0.5669) = 1 - P(Z \leq 0.5669) = 1 - 0.7155 = 0.2845$$

유의수준 5% 하에서 “보상 보험금액이 증가하지 않았다.” 라는 귀무가설을 기각할만한 근거가 충분하지 않다. 따라서 보상 보험금액이 지속적으로 증가하였다고 보기 어렵다.